

# Tail product-limit process for truncated data with application to extreme value index estimation

Souad Benchaira, Djamel Meraghni, Abdelhakim Necir\*

*Laboratory of Applied Mathematics, Mohamed Khider University, Biskra, Algeria*

## Abstract

A weighted Gaussian approximation to tail product-limit process for Pareto-like distributions of randomly right-truncated data is provided and a new consistent and asymptotically normal estimator of the extreme value index is derived. A simulation study is carried out to evaluate the finite sample behavior of the proposed estimator.

**Keywords:** Empirical process; Extreme value index; Heavy-tails; Hill estimator; Lynden-Bell estimator; Random truncation.

**AMS 2010 Subject Classification:** 60G70, 60F17, 62G30.

\*Corresponding author: [necirabdelhakim@yahoo.fr](mailto:necirabdelhakim@yahoo.fr)

*E-mail addresses:*

[benchaira.s@hotmail.fr](mailto:benchaira.s@hotmail.fr) (S. Benchaira)

[djmeraghni@yahoo.com](mailto:djmeraghni@yahoo.com) (D. Meraghni)

## 1. Introduction

Let  $(\mathbf{X}_i, \mathbf{Y}_i)$ ,  $1 \leq i \leq N$  be a sample of size  $N \geq 1$  from a couple  $(\mathbf{X}, \mathbf{Y})$  of independent random variables (rv's) defined over some probability space  $(\Omega, \mathcal{A}, \mathbf{P})$ , with continuous marginal distribution functions (df's)  $\mathbf{F}$  and  $\mathbf{G}$  respectively. Suppose that  $\mathbf{X}$  is truncated to the right by  $\mathbf{Y}$ , in the sense that  $\mathbf{X}_i$  is only observed when  $\mathbf{X}_i \leq \mathbf{Y}_i$ . This model of randomly truncated data commonly finds its applications in such areas like astronomy, economics, medicine and insurance. In the actuarial world, for instance, it is usual that the insurer claim data do not correspond to the underlying losses, because they are truncated from above. Indeed, when facing large claims, the insurance company specifies an upper limit to the amounts to be paid out. The excesses over this fixed threshold are then covered by a reinsurance company. This kind of reinsurance is called excess-loss reinsurance (see, e.g., [Embrechts et al., 1997](#)). In life insurance, the upper limit, which may be random, is called the cedent company retention level whereas in non-life insurance, it is called the deductible. The usefulness of the statistical analysis under random truncation is shown in [Herbst \(1999\)](#), where the author applies truncated model techniques to estimate loss reserves for IBNR (incurred but not reported) claim amounts. For a recent discussion on randomly right-truncated insurance claims, one refers to [Escudero and Ortega \(2008\)](#). Some examples of truncated data from astronomy and economics can be found in [Woodroffe \(1985\)](#) and for applications in the analysis of AIDS data, see [Wang \(1989\)](#). In reliability, a real dataset, consisting in lifetimes of automobile brake pads and already considered by [Lawless \(2002\)](#) in page 69, was recently analyzed in ? as an application of randomly truncated heavy-tailed models. Since the focus is on datasets that contain extreme values, then it would be natural to assume that both survival functions  $\overline{\mathbf{F}} := 1 - \mathbf{F}$  and  $\overline{\mathbf{G}} := 1 - \mathbf{G}$  are regularly varying at infinity with tail indices  $\gamma_1 > 0$  and  $\gamma_2 > 0$  respectively. That is, we have, for any  $x > 0$ ,

$$\lim_{z \rightarrow \infty} \frac{\overline{\mathbf{F}}(xz)}{\overline{\mathbf{F}}(z)} = x^{-1/\gamma_1} \text{ and } \lim_{z \rightarrow \infty} \frac{\overline{\mathbf{G}}(xz)}{\overline{\mathbf{G}}(z)} = x^{-1/\gamma_2}. \quad (1.1)$$

This class of distributions, which includes models such as Pareto, Burr, Fréchet, Lévy-stable and log-gamma, takes a prominent role in extreme value theory. Also known as heavy-tailed, Pareto-type or Pareto-like distributions, they provide appropriate descriptions for large insurance claims, log-returns, large price fluctuations, etc... (see, e.g., [Resnick, 2006](#)).

Let us denote  $(X_i, Y_i)$ ,  $i = 1, \dots, n$  to be the observed data, as copies of a couple of rv's  $(X, Y)$ , corresponding to the truncated sample  $(\mathbf{X}_i, \mathbf{Y}_i)$ ,  $i = 1, \dots, N$ , where  $n = n_N$  is a sequence of discrete rv's. By of the law of large numbers, we have  $n/N \xrightarrow{\mathbf{P}} p := \mathbf{P}(\mathbf{X} \leq \mathbf{Y})$  as  $n \xrightarrow{\mathbf{P}} \infty$ . For convenience, throughout the paper, the convergence in probability of  $n$  and/or any of its subsequences is simply denoted by  $\rightarrow$  instead of  $\xrightarrow{\mathbf{P}}$ . The joint distribution of  $X_i$  and  $Y_i$  is

$$\begin{aligned} H(x, y) &:= \mathbf{P}(X \leq x, Y \leq y) \\ &= \mathbf{P}(\mathbf{X} \leq x, \mathbf{Y} \leq y \mid \mathbf{X} \leq \mathbf{Y}) = p^{-1} \int_0^y \mathbf{F}(\min(x, z)) d\mathbf{G}(z). \end{aligned}$$

The marginal df's of the observed  $X$ 's and  $Y$ 's, respectively denoted by  $F$  and  $G$ , are equal to

$$F(x) := p^{-1} \int_0^x \overline{\mathbf{G}}(z) d\mathbf{F}(z) \quad \text{and} \quad G(y) := p^{-1} \int_0^y \mathbf{F}(z) d\mathbf{G}(z).$$

It follows that the corresponding tails are

$$\overline{F}(x) = -p^{-1} \int_x^\infty \overline{\mathbf{G}}(z) d\overline{\mathbf{F}}(z) \quad \text{and} \quad \overline{G}(y) = -p^{-1} \int_y^\infty \mathbf{F}(z) d\overline{\mathbf{G}}(z). \quad (1.2)$$

It is clear that the asymptotic behavior of  $\overline{F}$  simultaneously depends on  $\overline{\mathbf{G}}$  and  $\overline{\mathbf{F}}$ , while that of  $\overline{G}$  only relies on  $\overline{\mathbf{G}}$ . Making use of Proposition B.1.10 in [de Haan and Ferreira \(2006\)](#), for the regularly varying functions  $\overline{\mathbf{F}}$  and  $\overline{\mathbf{G}}$ , we may readily show that both  $\overline{G}$  and  $\overline{F}$  are regularly varying at infinity as well, with respective tail indices  $\gamma_2$  and  $\gamma := \gamma_1 \gamma_2 / (\gamma_1 + \gamma_2)$ . It is worth noting that the issue of analyzing extreme values in the context of random truncation, is at an early stage. Indeed, the first contribution was made in the recent paper of [Gardes and Stupfler \(2015\)](#), where the authors exploited the above relation between the three indices to define an estimator of  $\gamma_1$  by considering the classical Hill estimators of  $\gamma$  and  $\gamma_2$  ([Hill, 1975](#)) as functions of two distinct numbers of top statistics. However, they did not handle the case where these numbers are equal because of the difficulty in assessing the dependence between the two Hill estimators. In the present work, we introduce a tail product-limit process for which we provide a weighted Gaussian approximation as well. This tool will be very helpful when dealing with the estimation of any tail related quantity. In particular, it will lead to the asymptotic normality of the extreme value index estimator that we define, under random right-truncation, as a function of a single sample fraction of upper order statistics. But, prior to describing our estimation methodology, let us note that, as mentioned by [Gardes and Stupfler \(2015\)](#), in order to ensure that it remains enough extreme data for the inference to be accurate,

we need to impose the condition  $\gamma_1 < \gamma_2$ . In other words, we consider the situation where the tail of the rv of interest  $\mathbf{X}$  is not too contaminated by the truncation rv  $\mathbf{Y}$ . Since  $\mathbf{F}$  and  $\mathbf{G}$  are heavy-tailed, then their right endpoints are infinite and thus they are equal. Hence, from [Woodroffe \(1985\)](#), we may write

$$\int_x^\infty \frac{d\mathbf{F}(y)}{\mathbf{F}(y)} = \int_x^\infty \frac{dF(y)}{C(y)}, \quad (1.3)$$

where

$$C(z) := \mathbf{P}(X \leq z \leq Y) = F(z) - G(z). \quad (1.4)$$

Differentiating (1.3) leads to the following crucial equation

$$C(x) d\mathbf{F}(x) = \mathbf{F}(x) dF(x), \quad (1.5)$$

(see, for instance, [Strzalkowska-Kominiak and Stute, 2009](#)), whose solution is defined by  $\mathbf{F}(x) = \exp - \int_x^\infty dF(z)/C(z)$ . Replacing  $F$  and  $C$  by their respective empirical counterparts

$$F_n(x) := n^{-1} \sum_{i=1}^n \mathbf{1}(X_i \leq x) \text{ and } C_n(x) := n^{-1} \sum_{i=1}^n \mathbf{1}(X_i \leq x \leq Y_i),$$

yields the product-limit estimator

$$\mathbf{F}_n(x) := \prod_{i: X_i > x} \exp \left\{ -\frac{1}{nC_n(X_i)} \right\},$$

to the underlying df  $\mathbf{F}$ . The first mathematical investigation on this estimator may be attributed to [Woodroffe \(1985\)](#) and the central limit theorem under random truncation was established by [Stute and Wang \(2008\)](#). Note that the approximation  $\exp(-t) \sim 1 - t$ , for small  $t > 0$ , results in the well-known estimator introduced by [Lynden-Bell \(1971\)](#). Let us now introduce a tail product-limit process corresponding to  $\mathbf{F}_n$  as follows:

$$\mathbf{D}_n(x) := \sqrt{k} \left( \frac{\overline{\mathbf{F}}_n(xX_{n-k:n})}{\overline{\mathbf{F}}_n(X_{n-k:n})} - x^{-1/\gamma_1} \right), \quad x > 0, \quad (1.6)$$

where  $X_{1:n} \leq \dots \leq X_{n:n}$  denote the order statistics pertaining to  $X_1, \dots, X_n$  and  $k = k_n$  is a sequence of discrete rv's satisfying

$$1 < k < n, \quad k \rightarrow \infty \text{ and } k/n \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (1.7)$$

Observe that, in the case of complete data we have  $\mathbf{F}_n \equiv F_n$  with  $\overline{F}_n(X_{n-k:n}) = k/n$  and thus the process defined in (1.6) becomes

$$D_n(x) := \sqrt{k} \left( \frac{n}{k} \overline{F}_n(xX_{n-k:n}) - x^{-1/\gamma_1} \right).$$

By jointly applying Theorems 2.4.8 and 5.1.4 (pages 52 and 161) in [de Haan and Ferreira \(2006\)](#) we have that, for  $x_0 > 0$  and  $0 < \xi < 1/2$ ,

$$\sup_{x \geq x_0} x^{(1/2-\xi)/\gamma_1} \left| D_n(x) - \Gamma(x; W) - x^{-1/\gamma_1} \frac{x^{\tau_1/\gamma_1} - 1}{\tau_1 \gamma_1} \sqrt{k} A_0(n/k) \right| \xrightarrow{\mathbf{P}} 0, \quad (1.8)$$

provided that  $F$  fulfills the second-order regular variation condition with auxiliary function  $A_0$  tending to zero, not changing sign near infinity, having a regularly varying absolute value with index  $\tau_1 < 0$  and satisfying  $\sqrt{k} A_0(n/k) = O(1)$ . Here  $\Gamma(x; W) := W(x^{-1/\gamma_1}) - x^{-1/\gamma_1} W(1)$  with  $\{W(s); 0 \leq s \leq 1\}$  being a standard Wiener process. Many authors used this approximation to establish the limit distributions of several statistics of heavy-tailed distributions, such as tail index estimators (see, e.g., [de Haan and Ferreira, 2006](#), page 76) and goodness-of-fit statistics ([Koning and Peng, 2008](#)). The main goal of this paper is to provide an analogous result to (1.8) in the random truncation setting through the tail product-limit process (1.6), which, to the best of our knowledge, was not addressed yet in the extreme value theory literature.

The rest of the paper is organized as follows. In Section 2, we present our main result which consists in a Gaussian approximation to the tail product-limit process  $\mathbf{D}_n(x)$ . As an application, we introduce, in Section 3, a new Hill-type estimator ([Hill, 1975](#)) for the tail index  $\gamma_1$  and we establish its consistency and asymptotic normality. The finite sample behavior of the proposed estimator is checked by simulation in Section 4. The proofs are postponed to Section 5 and some results that are instrumental to our needs are gathered in two lemmas in the Appendix.

## 2. Main results

Weak approximations of extreme value theory based statistics are achieved in the second-order framework (see [de Haan and Stadtmüller, 1996](#)). Thus, it seems quite natural to suppose that df's  $\mathbf{F}$  and  $\mathbf{G}$  satisfy the well-known second-order condition of regular variation that we express in terms of the tail quantile functions. That is, we assume that for  $x > 0$ , we have

$$\lim_{t \rightarrow \infty} \frac{\mathbb{U}_{\mathbf{F}}(tx)/\mathbb{U}_{\mathbf{F}}(t) - x^{\gamma_1}}{\mathbf{A}_{\mathbf{F}}(t)} = x^{\gamma_1} \frac{x^{\tau_1} - 1}{\tau_1}, \quad (2.9)$$

and

$$\lim_{t \rightarrow \infty} \frac{\mathbb{U}_{\mathbf{G}}(tx)/\mathbb{U}_{\mathbf{G}}(t) - x^{\gamma_2}}{\mathbf{A}_{\mathbf{G}}(t)} = x^{\gamma_2} \frac{x^{\tau_2} - 1}{\tau_2}, \quad (2.10)$$

where  $\tau_1, \tau_2 < 0$  are the second-order parameters and  $\mathbf{A}_{\mathbf{F}}, \mathbf{A}_{\mathbf{G}}$  are functions tending to zero and not changing signs near infinity with regularly varying absolute values

at infinity with indices  $\tau_1, \tau_2$  respectively. For any df  $K$ , the function  $\mathbb{U}_K(t) := K^\leftarrow(1 - 1/t)$ ,  $t > 1$ , stands for the tail quantile function.

**Theorem 2.1.** *Assume that both second-order conditions (2.9) and (2.10) hold with  $\gamma_1 < \gamma_2$ . Let  $k = k_n$  be a sequence satisfying (1.7), then there exist a function  $\mathbf{A}_0(t) \sim \mathbf{A}_F(1/\bar{\mathbf{F}}(\mathbb{U}_F(t)))$  and a standard Wiener process  $\{\mathbf{W}(s); 0 \leq s \leq 1\}$ , defined on the probability space  $(\Omega, \mathcal{A}, \mathbf{P})$ , such that, for  $0 < \xi < 1/2 - \gamma/\gamma_2$  and  $x_0 > 0$ , we have*

$$\sup_{x \geq x_0} x^{(1/2-\xi)/\gamma-1/\gamma_2} \left| \mathbf{D}_n(x) - \Gamma(x; \mathbf{W}) - x^{-1/\gamma_1} \frac{x^{\tau_1/\gamma_1} - 1}{\gamma_1 \tau_1} \sqrt{k} \mathbf{A}_0(n/k) \right| \xrightarrow{\mathbf{P}} 0,$$

as  $n \rightarrow \infty$ , provided that  $\sqrt{k} \mathbf{A}_0(n/k) = O(1)$ , where  $\{\Gamma(x; \mathbf{W}); x > 0\}$  is a Gaussian process defined by

$$\begin{aligned} \Gamma(x; \mathbf{W}) &:= \frac{\gamma}{\gamma_1} x^{-1/\gamma_1} \{x^{1/\gamma} \mathbf{W}(x^{-1/\gamma}) - \mathbf{W}(1)\} \\ &+ \frac{\gamma}{\gamma_1 + \gamma_2} x^{-1/\gamma_1} \int_0^1 s^{-\gamma/\gamma_2-1} \{x^{1/\gamma} \mathbf{W}(x^{-1/\gamma}s) - \mathbf{W}(s)\} ds. \end{aligned}$$

**Remark 2.1.** *A very large value of  $\gamma_2$  yields a  $\gamma$ -value that is very close to  $\gamma_1$ , meaning that the really observed sample is almost the whole dataset. In other words, the complete data case corresponds to the situation when  $1/\gamma_2 \equiv 0$ , in which case we have  $\gamma \equiv \gamma_1$ . It follows that*

$$\frac{\gamma}{\gamma_1 + \gamma_2} \int_0^1 s^{-\gamma/\gamma_2-1} \{x^{1/\gamma} \mathbf{W}(x^{-1/\gamma}s) - \mathbf{W}(s)\} ds \equiv 0,$$

and therefore  $\Gamma(x; \mathbf{W}) = \mathbf{W}(x^{-1/\gamma_1}) - x^{-1/\gamma_1} \mathbf{W}(1)$ , which agrees with the weak approximation (1.8).

### 3. Tail index estimation

We start the construction of our estimator by noting that from Theorem 1.2.2 in [de Haan and Ferreira \(2006\)](#), the first-order condition (1.1) (for  $\bar{\mathbf{F}}$ ) implies that

$$\lim_{t \rightarrow \infty} \frac{1}{\bar{\mathbf{F}}(t)} \int_t^\infty x^{-1} \bar{\mathbf{F}}(x) dx = \gamma_1,$$

which, by an integration by parts, becomes

$$\lim_{t \rightarrow \infty} \frac{1}{\bar{\mathbf{F}}(t)} \int_t^\infty \log \frac{x}{t} d\mathbf{F}(x) = \gamma_1. \quad (3.11)$$

Replacing  $\mathbf{F}$  by  $\mathbf{F}_n$  and letting  $t = X_{n-k:n}$  yields

$$\hat{\gamma}_1 := \frac{1}{\bar{\mathbf{F}}_n(X_{n-k:n})} \int_{X_{n-k:n}}^\infty \log \frac{x}{X_{n-k:n}} d\mathbf{F}_n(x),$$

as a new estimator to  $\gamma_1$ . By setting  $\varphi_n^{(1)}(x) := \mathbf{1}\{x \geq X_{n-k:n}\} \log(x/X_{n-k:n})$  and  $\varphi_n^{(2)}(x) := \mathbf{1}\{x \geq X_{n-k:n}\}$ , this may be rewritten into

$$\hat{\gamma}_1 = \frac{\int_0^\infty \varphi_n^{(1)}(x) d\mathbf{F}_n(x)}{\int_0^\infty \varphi_n^{(2)}(x) d\mathbf{F}_n(x)}.$$

From the empirical counterpart of equation (1.5) we get

$$\int_0^\infty \varphi_n^{(1)}(x) d\mathbf{F}_n(x) = \frac{1}{n} \sum_{i=n-k}^n \frac{\mathbf{F}_n(X_{i:n})}{C_n(X_{i:n})} \log(X_{i:n}/X_{n-k:n}),$$

and

$$\int_0^\infty \varphi_n^{(2)}(x) d\mathbf{F}_n(x) = \frac{1}{n} \sum_{i=n-k}^n \frac{\mathbf{F}_n(X_{i:n})}{C_n(X_{i:n})}.$$

Finally, changing  $i$  to  $n - i + 1$  yields

$$\hat{\gamma}_1 = \left( \sum_{i=1}^k \frac{\mathbf{F}_n(X_{n-i+1:n})}{C_n(X_{n-i+1:n})} \right)^{-1} \sum_{i=1}^k \frac{\mathbf{F}_n(X_{n-i+1:n})}{C_n(X_{n-i+1:n})} \log \frac{X_{n-i+1:n}}{X_{n-k:n}}.$$

**Remark 3.1.** For complete data, we have  $\mathbf{F}_n \equiv F_n \equiv C_n$  and consequently  $\hat{\gamma}_1$  reduces to the classical Hill estimator (Hill, 1975).

**Theorem 3.1.** Assume that (1.1) holds with  $\gamma_1 < \gamma_2$  and let  $k = k_n$  be an integer sequence satisfying (1.7). Then  $\hat{\gamma}_1 \rightarrow \gamma_1$  in probability. Assume further that both second-order conditions (2.9) and (2.10) hold and  $\sqrt{k}\mathbf{A}_0(n/k) = O(1)$ , then

$$\begin{aligned} \sqrt{k}(\hat{\gamma}_1 - \gamma_1) &= \frac{\sqrt{k}\mathbf{A}_0(n/k)}{1 - \tau_1} - \gamma \mathbf{W}(1) \\ &+ \frac{\gamma}{\gamma_1 + \gamma_2} \int_0^1 (\gamma_2 - \gamma_1 - \gamma \log s) s^{-\gamma/\gamma_2 - 1} \mathbf{W}(s) ds + o_{\mathbf{P}}(1). \end{aligned}$$

**Corollary 3.1.** If, in addition to the assumptions of Theorem 3.1, we suppose that  $\sqrt{k}\mathbf{A}_0(n/k) \rightarrow \lambda$ , then

$$\sqrt{k}(\hat{\gamma}_1 - \gamma_1) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\lambda}{1 - \tau_1}, \sigma^2\right), \text{ as } n \rightarrow \infty,$$

where

$$\sigma^2 := \gamma^2 (1 + \gamma_1/\gamma_2) (1 + (\gamma_1/\gamma_2)^2) / (1 - \gamma_1/\gamma_2)^3.$$

**Remark 3.2.** In the case of complete data we have, from Remark 2.1,  $\sigma^2 \equiv \gamma_1^2$ . It follows that  $\sqrt{k}(\hat{\gamma}_1 - \gamma_1) \xrightarrow{\mathcal{D}} \mathcal{N}(\lambda/(1 - \tau_1), \gamma_1^2)$ , as  $n \rightarrow \infty$ , which meets the asymptotic normality of the classical Hill estimator (Hill, 1975), see for instance, Theorem 3.2.5 in de Haan and Ferreira (2006).

#### 4. Simulation study

This study, just intended for illustrating the performance of our estimator, is realized through two sets of truncated and truncation data, both drawn from Burr's model:

$$\overline{\mathbf{F}}(x) = (1 + x^{1/\delta})^{-\delta/\gamma_1}, \quad \overline{\mathbf{G}}(x) = (1 + x^{1/\delta})^{-\delta/\gamma_2}, \quad x \geq 0,$$

where  $\delta, \gamma_1, \gamma_2 > 0$ . The corresponding percentage of observed data is equal to  $p = \gamma_2/(\gamma_1 + \gamma_2)$ . We fix  $\delta = 1/4$  and choose the values 0.6 and 0.8 for  $\gamma_1$  and 70%, 80% and 90% for  $p$ . For each couple  $(\gamma_1, p)$ , we solve the equation  $p = \gamma_2/(\gamma_1 + \gamma_2)$  to get the pertaining  $\gamma_2$ -value. We vary the common size  $N$  of both samples  $(\mathbf{X}_1, \dots, \mathbf{X}_N)$  and  $(\mathbf{Y}_1, \dots, \mathbf{Y}_N)$ , then for each size, we generate 1000 independent replicates. Our overall results are taken as the empirical means of the results obtained through all repetitions. To determine the optimal number (that we denote by  $k^*$ ) of upper order statistics used in the computation of  $\widehat{\gamma}_1$ , we apply the algorithm of (Reiss and Thomas, 2007, page 137). The performance of the newly defined estimator, in terms of absolute bias and root of the mean squared error (rmse) is summarized in Table 4.1, where we see that, as expected, the size of the initial sample influences the estimation: the larger  $N$ , the better the estimation. On the other hand, we note that the estimation accuracy decreases when the truncation percentage increases, which seems logical. Finally, we observe that the estimation of the larger value of the tail index is less precise.

#### 5. Proofs

**5.1. Proof of Theorem 2.1.** Set  $U_i := \overline{F}(X_i)$  and define the corresponding uniform tail empirical process by  $\alpha_n(s) := \sqrt{k}(\mathbf{U}_n(s) - s)$ , for  $0 \leq s \leq 1$ , where  $\mathbf{U}_n(s) := k^{-1} \sum_{i=1}^n \mathbf{1}(U_i < ks/n)$ . The weighted weak approximation to  $\alpha_n(s)$  given in terms of, either a sequence of Wiener processes (see, e.g., Einmahl, 1992 and Dress et al., 2006) or a single Wiener process as in Proposition 3.1 of Einmahl et al. (2006), will be very crucial to our proof procedure. In the sequel, we use the latter representation which says that: there exists a Wiener process  $\mathbf{W}$ , such that for every  $0 \leq \eta < 1/2$ ,

$$\sup_{0 \leq s \leq 1} s^{-\eta} |\alpha_n(s) - \mathbf{W}(s)| \xrightarrow{\mathbf{P}} 0, \quad \text{as } n \rightarrow \infty. \quad (5.12)$$

We begin by fixing  $x_0 > 0$ , then we decompose  $k^{-1/2} \mathbf{D}_n(x)$ , for  $x \geq x_0$ , as the sum of the following four terms:

$$\mathbf{M}_{n1}(x) := x^{-1/\gamma_1} \frac{\overline{\mathbf{F}}_n(xX_{n-k:n}) - \overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}(xX_{n-k:n})},$$



	$p = 0.7$							
	$\gamma_1 = 0.6$				$\gamma_1 = 0.8$			
$N$	$n$	$k^*$	absolute bias	rmse	$n$	$k^*$	absolute bias	rmse
200	139	8	0.1811	0.4645	140	9	0.2485	0.6034
300	210	17	0.1280	0.3451	209	23	0.1230	0.5082
500	348	28	0.1151	0.2803	348	31	0.1024	0.3708
1000	699	43	0.0461	0.2421	700	42	0.0684	0.3825
1500	1050	67	0.0212	0.2362	1050	57	0.0527	0.2539
2000	1399	95	0.0254	0.2261	1401	94	0.0469	0.2602
	$p = 0.8$							
200	159	10	0.0759	0.5235	159	11	0.1435	0.6815
300	240	21	0.0485	0.3647	238	18	0.1052	0.4511
500	399	34	0.0403	0.2718	400	41	0.0935	0.3170
1000	800	55	0.0408	0.1999	800	54	0.0504	0.2881
1500	1205	116	0.0562	0.1911	1199	85	0.0428	0.2362
2000	1599	117	0.0285	0.1534	1599	117	0.0373	0.2054
	$p = 0.9$							
200	180	14	0.0537	0.5204	179	17	0.1098	0.7531
300	269	24	0.0388	0.3343	268	26	0.0844	0.3654
500	450	48	0.0294	0.2869	450	49	0.0721	0.2448
1000	899	64	0.0359	0.1557	899	67	0.0465	0.2014
1500	1349	103	0.0171	0.1350	1349	101	0.0385	0.1828
2000	1799	144	0.0188	0.1107	1799	145	0.0251	0.1466

TABLE 4.1. Absolute bias and rmse of the tail index estimator based on 1000 right-truncated samples of Burr models.

$$\mathbf{M}_{n2}(x) := -\frac{\overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}_n(X_{n-k:n})} \frac{\overline{\mathbf{F}}_n(X_{n-k:n}) - \overline{\mathbf{F}}(X_{n-k:n})}{\overline{\mathbf{F}}(X_{n-k:n})},$$

$$\mathbf{M}_{n3}(x) := \left( \frac{\overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}_n(X_{n-k:n})} - x^{-1/\gamma_1} \right) \frac{\overline{\mathbf{F}}_n(xX_{n-k:n}) - \overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}(xX_{n-k:n})}$$

and

$$\mathbf{M}_{n4}(x) := \frac{\overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}(X_{n-k:n})} - x^{-1/\gamma_1}.$$

In order to establish the result of the theorem, we will successively show that, under the first-order of regular variation conditions, we have uniformly on  $x \geq x_0$ , for  $\gamma/\gamma_2 < \eta < 1/2$  and  $\epsilon > 0$  sufficiently small

$$\begin{aligned} & x^{1/\gamma_1} \sqrt{k} \mathbf{M}_{n1}(x) \\ &= x^{1/\gamma} \left\{ \frac{\gamma}{\gamma_1} \mathbf{W}(x^{-1/\gamma}) + \frac{\gamma}{\gamma_1 + \gamma_2} \int_0^1 t^{-\gamma/\gamma_2 - 1} \mathbf{W}(x^{-1/\gamma} t) dt \right\} + O_{\mathbf{P}}(\epsilon) x^{(1-\eta)/\gamma \pm \epsilon}, \\ & x^{1/\gamma_1} \sqrt{k} \mathbf{M}_{n2}(x) = - \left\{ \frac{\gamma}{\gamma_1} \mathbf{W}(1) + \frac{\gamma}{\gamma_1 + \gamma_2} \int_0^1 t^{-\gamma/\gamma_2 - 1} \mathbf{W}(t) dt \right\} + O_{\mathbf{P}}(\epsilon) x^{\pm \epsilon}, \end{aligned}$$

and

$$x^{1/\gamma_1} \sqrt{k} \mathbf{M}_{n3}(x) = O_{\mathbf{P}}(\epsilon) x^{-1/\gamma_1 + (1-\eta)/\gamma \pm \epsilon}.$$

Moreover, if we assume the second-order condition we will show that

$$x^{1/\gamma_1} \sqrt{k} \mathbf{M}_{n4}(x) = (1 + o_{\mathbf{P}}(1)) \frac{x^{\tau_1/\gamma_1} - 1}{\gamma_1 \tau_1} \sqrt{k} \mathbf{A}_0(n/k).$$

Here  $O_{\mathbf{P}}$  and  $o_{\mathbf{P}}$  stand for the usual stochastic order symbols. For convenience, let  $a_k := \mathbb{U}_F(n/k)$  and recall that  $\mathbb{U}_F$  is regularly varying (with index  $\gamma$ ). Then by combining Corollary 2.2.2 with Proposition B.1.10 in [de Haan and Ferreira \(2006\)](#), we show that  $X_{n-k:n}/a_k \xrightarrow{\mathbf{P}} 1$  as  $n \rightarrow \infty$ , which implies, due to the regular variation of  $\overline{\mathbf{F}}$ , that  $\overline{\mathbf{F}}(xa_k)/\overline{\mathbf{F}}(xX_{n-k:n}) = 1 + o_{\mathbf{P}}(1)$  and therefore

$$\mathbf{M}_{n1}(x) = (1 + o_{\mathbf{P}}(1)) \mathbf{M}_{n1}^*(x), \quad (5.13)$$

where

$$\mathbf{M}_{n1}^*(x) := x^{-1/\gamma_1} \frac{\overline{\mathbf{F}}_n(xX_{n-k:n}) - \overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}(xa_k)}.$$

Now, observe that, in view of equation (1.5), we may write

$$\mathbf{F}(x) = \exp -\Lambda(x) \text{ and } \mathbf{F}_n(x) = \exp -\Lambda_n(x),$$

where  $\Lambda(x)$  and its empirical counterpart  $\Lambda_n(x)$  are defined by  $\int_x^\infty dF(z)/C(z)$  and  $\int_x^\infty dF_n(z)/C_n(z)$  respectively. Note that  $\overline{\mathbf{F}}_n(xX_{n-k:n})$ ,  $\overline{\mathbf{F}}(xX_{n-k:n})$  and  $\overline{\mathbf{F}}(xa_k)$  tend to zero in probability, uniformly on  $x \geq x_0$ , it follows that  $\Lambda_n(xX_{n-k:n})$ ,  $\Lambda(xX_{n-k:n})$  and  $\Lambda(xa_k)$  go to zero in probability as well. Using the approximation  $1 - \exp(-t) \sim t$ , as  $t \downarrow 0$ , we may write

$$x^{1/\gamma_1} \mathbf{M}_{n1}^*(x) = (1 + o_{\mathbf{P}}(1)) \frac{\Lambda_n(xX_{n-k:n}) - \Lambda(xX_{n-k:n})}{\Lambda(xa_k)}.$$

Next, we provide a Gaussian approximation to the expression

$$\sqrt{k} \frac{\Lambda_n(xX_{n-k:n}) - \Lambda(xX_{n-k:n})}{\Lambda(xa_k)},$$

then we deduce one to  $\sqrt{k}x^{1/\gamma_1}\mathbf{M}_{n1}^*(x)$ . For this, we decompose the difference  $\Lambda_n(xX_{n-k:n}) - \Lambda(xX_{n-k:n})$  in the sum of

$$S_{n1}(x) := - \int_{xa_k}^{\infty} \frac{d(\overline{F}_n(z) - \overline{F}(z))}{C(z)},$$

$$S_{n2}(x) := - \int_{xX_{n-k:n}}^{\infty} \left\{ \frac{1}{C_n(z)} - \frac{1}{C(z)} \right\} d\overline{F}_n(z),$$

and

$$S_{n3}(x) := - \int_{xX_{n-k:n}}^{xa_k} \frac{d(\overline{F}_n(z) - \overline{F}(z))}{C(z)}.$$

For the first term, we use the fact that  $\overline{F}_n(z) = 0$  for  $z \geq X_{n:n}$ , to write, after an integration by parts and a change of variables,  $S_{n1}(x) = S_{n1}^{(1)}(x) - S_{n1}^{(2)}(x)$ , with

$$S_{n1}^{(1)}(x) := \frac{\overline{F}_n(a_kx) - \overline{F}(a_kx)}{C(a_kx)} \text{ and } S_{n1}^{(2)}(x) := \int_x^{\infty} \frac{\overline{F}_n(a_kz) - \overline{F}(a_kz)}{C^2(a_kz)} dC(a_kz).$$

It is easy to verify that  $\overline{F}_n(xa_k) - \overline{F}(xa_k) = \frac{\sqrt{k}}{n} \alpha_n \left( \frac{n}{k} \overline{F}(xa_k) \right)$ , it follows that

$$\frac{\sqrt{k}S_{n1}^{(1)}(x)}{\Lambda(a_kx)} = d_n(x) \alpha_n \left( \frac{n}{k} \overline{F}(a_kx) \right),$$

where  $d_n(x) := \frac{k/n}{\Lambda(xa_k)C(a_kx)}$ . From Lemma 6.2 (iii), we have

$$d_n(x) = (\gamma/\gamma_1) x^{1/\gamma} + O(\epsilon) x^{1/\gamma \pm \epsilon}, \quad (5.14)$$

as  $n \rightarrow \infty$ , uniformly on  $x \geq x_0$ , it follows that

$$\frac{\sqrt{k}S_{n1}^{(1)}(x)}{\Lambda(a_kx)} = \{(\gamma/\gamma_1) x^{1/\gamma} + O_{\mathbf{P}}(\epsilon) x^{1/\gamma \pm \epsilon}\} \alpha_n \left( \frac{n}{k} \overline{F}(a_kx) \right).$$

On the other hand, for  $0 < \eta < 1/2$ , the sequence of rv's  $\sup_{0 < s \leq 1} |\alpha_n(s)|/s^\eta$  is stochastically bounded. This comes from the inequality

$$\sup_{0 < s \leq 1} s^{-\eta} |\alpha_n(s)| \leq \sup_{0 < s \leq 1} s^{-\eta} |\alpha_n(s) - \mathbf{W}(s)| + \sup_{0 < s \leq 1} s^{-\eta} |\mathbf{W}(s)|,$$

with approximation (5.12) and the fact  $\sup_{0 < s \leq 1} s^{-\eta} |\mathbf{W}(s)| = O_{\mathbf{P}}(1)$  (see, e.g., Lemma 3.2 in Einmahl et al., 2006). Now, let  $\epsilon > 0$  be sufficiently small. Then, by applying Potter's inequalities to  $\overline{F}$  (see, e.g., Proposition B.1.9, assertion 5 in de Haan and Ferreira, 2006), we write  $\frac{n}{k} \overline{F}(a_kx) \leq (1 + \epsilon) x^{-1/\gamma \pm \epsilon}$ , it follows that  $\alpha_n \left( \frac{n}{k} \overline{F}(a_kx) \right) = O_{\mathbf{P}}(x^{-\eta/\gamma \pm \eta\epsilon})$ . For notational simplicity and without loss of generality, we attribute  $\epsilon$  to any constant times  $\epsilon$  and  $v^{\pm\epsilon}$  to any linear combinations of  $v^{\pm c_1\epsilon}$  and  $v^{\pm c_2\epsilon}$ , for every  $c_1, c_2 > 0$ . Therefore

$$\frac{\sqrt{k}S_{n1}^{(1)}(x)}{\Lambda(a_kx)} = \frac{\gamma}{\gamma_1} x^{1/\gamma} \alpha_n \left( \frac{n}{k} \overline{F}(a_kx) \right) + O_{\mathbf{P}}(\epsilon) x^{(1-\eta)/\gamma \pm \epsilon}.$$

For  $S_{n1}^{(2)}(x)$ , let us write

$$\frac{\sqrt{k}S_{n1}^{(2)}(x)}{\Lambda(a_k x)} = d_n(x) \frac{C(a_k x)}{C(a_k)} \int_x^\infty \frac{C^2(a_k)}{C^2(a_k z)} \alpha_n\left(\frac{n}{k}\overline{F}(a_k z)\right) d\frac{C(a_k z)}{C(a_k)}.$$

From Lemma 6.2 (i), the function  $C$  is regularly varying at infinity with index  $(-1/\gamma_2)$ , as  $\overline{G}$  is, this implies that  $C(xa_k)/C(a_k) = x^{-1/\gamma_2} + O(\epsilon)x^{-1/\gamma_2 \pm \epsilon}$ . Then by using (5.14), we get

$$d_n(x) \frac{C(a_k x)}{C(a_k)} = (\gamma/\gamma_1) x^{1/\gamma_1} + O(\epsilon) x^{1/\gamma_1 \pm \epsilon}, \text{ as } n \rightarrow \infty. \quad (5.15)$$

For convenience, we set  $\sqrt{k}S_{n1}^{(2)}(x)/\Lambda(a_k x) = \{1 + O(\epsilon)x^{\pm\epsilon}\} \mathcal{T}_n(x)$ , where

$$\mathcal{T}_n(x) := \frac{\gamma}{\gamma_1} x^{1/\gamma_1} \int_x^\infty \frac{C^2(a_k)}{C^2(a_k z)} \alpha_n\left(\frac{n}{k}\overline{F}(a_k z)\right) d\frac{C(a_k z)}{C(a_k)},$$

which we decompose in the sum of

$$I_n(x) := \frac{\gamma}{\gamma_1} x^{1/\gamma_1} \int_x^\infty \frac{C^2(a_k)}{C^2(a_k z)} \alpha_n\left(\frac{n}{k}\overline{F}(a_k z)\right) d\frac{F(a_k z)}{C(a_k)},$$

$$J_n(x) := -\frac{\gamma}{\gamma_1} x^{1/\gamma_1} \int_x^\infty \left\{ \frac{C^2(a_k)}{C^2(a_k z)} - z^{2/\gamma_2} \right\} \alpha_n\left(\frac{n}{k}\overline{F}(a_k z)\right) d\frac{G(a_k z)}{C(a_k)},$$

and

$$K_n(x) := \frac{\gamma}{\gamma_1} x^{1/\gamma_1} \int_x^\infty z^{2/\gamma_2} \alpha_n\left(\frac{n}{k}\overline{F}(a_k z)\right) d\frac{\overline{G}(a_k z)}{C(a_k)}.$$

Recall that  $a_k \rightarrow \infty$ ,  $C(a_k) \sim \overline{G}(a_k)$  and  $\overline{F}(a_k) = o(\overline{G}(a_k))$  as  $n \rightarrow \infty$ . On the other hand, by using, once again, Potter's inequalities to  $C$ , (regularly varying at infinity with index  $-1/\gamma_2$ ), we write, for all large  $n$  and  $z \geq x$ ,

$$(1 - \epsilon) z^{-1/\gamma_2} \min(z^\epsilon, z^{-\epsilon}) \leq \frac{C(a_k z)}{C(a_k)} \leq (1 + \epsilon) z^{-1/\gamma_2} \max(z^\epsilon, z^{-\epsilon}). \quad (5.16)$$

It is clear this implies that  $C^2(a_k)/C^2(a_k z) \leq (1 - \epsilon)^{-2} z^{2/\gamma_2 \pm 2\epsilon}$ . In view of the stochastic boundedness of  $\sup_{0 < s \leq 1} |\alpha_n(s)|/s^\eta$  and the fact that  $\frac{n}{k}\overline{F}(a_k z) \leq (1 + \epsilon) z^{-1/\gamma \pm \epsilon}$ , we have

$$I_n(x) = o_{\mathbf{P}}(1) x^{1/\gamma_1} \int_x^\infty z^{2/\gamma_2 \mp 2\epsilon} (z^{-1/\gamma \pm \epsilon})^\eta d\frac{\overline{F}(a_k z)}{\overline{F}(a_k)}.$$

Integrating by parts, we readily get  $I_n(x) = o_{\mathbf{P}}(1) x^{1/\gamma_2 - \eta/\gamma \pm \epsilon} = o_{\mathbf{P}}(1) x^{(1-\eta)/\gamma \pm \epsilon}$ . Let us now consider  $J_n(x)$ . From Proposition B.1.10 in [de Haan and Ferreira \(2006\)](#), we have  $|C(a_k z)/C(a_k) - z^{-1/\gamma_2}| \leq \epsilon z^{-1/\gamma_2 \pm \epsilon}$ . Applying the mean value theorem, then combining this inequality with (5.16), yield

$$\left| \frac{C^2(a_k)}{C^2(a_k z)} - z^{2/\gamma_2} \right| \leq \epsilon \frac{2(z^{\pm\epsilon} + 1)}{(1 - \epsilon)^3} z^{2/\gamma_2 \pm \epsilon}.$$

Similar arguments as the above lead to  $J_n(x) = O_{\mathbf{p}}(\epsilon) x^{(1-\eta)/\gamma \pm \epsilon}$ . Now, we focus on  $K_n(x)$ . Since  $C(a_k) \sim \overline{G}(a_k)$ , then

$$K_n(x) = (1 + o_{\mathbf{p}}(1)) \frac{\gamma}{\gamma_1} x^{1/\gamma_1} \int_x^\infty z^{2/\gamma_2} \alpha_n\left(\frac{n}{k} \overline{F}(a_k z)\right) d \frac{\overline{G}(a_k z)}{\overline{G}(a_k)}.$$

Let  $G^\leftarrow$  denote the quantile function pertaining to df  $G$  and use the change of variables  $z = G^\leftarrow(1 - s \overline{G}(a_k)) / a_k$  to get

$$K_n(x) = -(1 + o_{\mathbf{p}}(1)) \frac{\gamma}{\gamma_1} x^{1/\gamma_1} \int_0^{\frac{\overline{G}(a_k x)}{\overline{G}(a_k)}} \left( \frac{G^\leftarrow(1 - s \overline{G}(a_k))}{a_k} \right)^{2/\gamma_2} \alpha_n(\ell_n(s)) ds,$$

where  $\ell_n(s) := \frac{n}{k} \overline{F}(G^\leftarrow(1 - s \overline{G}(a_k)))$ . It is easy to check that

$$K_n(x) = -(1 + o_{\mathbf{p}}(1)) \sum_{i=1}^3 K_{ni}(x),$$

where

$$K_{n1}(x) := \frac{\gamma}{\gamma_1} x^{\frac{1}{\gamma_1}} \int_0^{\frac{\overline{G}(a_k x)}{\overline{G}(a_k)}} \left\{ \left( \frac{G^\leftarrow(1 - s \overline{G}(a_k))}{a_k} \right)^{2/\gamma_2} - s^{-2} \right\} \alpha_n(\ell_n(s)) ds,$$

$$K_{n2}(x) := \frac{\gamma}{\gamma_1} x^{1/\gamma_1} \int_{x^{-1/\gamma_2}}^{\frac{\overline{G}(a_k x)}{\overline{G}(a_k)}} s^{-2} \alpha_n(\ell_n(s)) ds,$$

and

$$K_{n3}(x) := \frac{\gamma}{\gamma_1} x^{1/\gamma_1} \int_0^{x^{-1/\gamma_2}} s^{-2} \alpha_n(\ell_n(s)) ds.$$

By routine manipulations and similar arguments based on stochastic boundedness of  $\sup_{0 < s \leq 1} |\alpha_n(s)| / s^\eta$  and the aforementioned Proposition B.1.10 applied to the regularly varying functions  $\overline{G}$  and  $G^\leftarrow(1 - \cdot)$ , we show that  $K_{ni}(x) = O_{\mathbf{p}}(\epsilon) x^{(1-\eta)/\gamma \mp \epsilon}$ ,  $i = 1, 2$  and  $K_{n3}(x) = O_{\mathbf{p}}(1) x^{(1-\eta)/\gamma \mp \epsilon}$ , therefore we omit the details. Up to this stage, we have shown that  $\mathcal{T}_n(x) = O_{\mathbf{p}}(1) x^{(1-\eta)/\gamma \mp \epsilon}$ . It follows that

$$\frac{\sqrt{k} S_{n1}^{(2)}(x)}{\Lambda(a_k x)} = \mathcal{T}_n(x) + O_{\mathbf{p}}(\epsilon) x^{(1-\eta)/\gamma \mp \epsilon},$$

which, after gathering the components of  $\mathcal{T}_n(x)$ , is equal to

$$\frac{\gamma}{\gamma_1} x^{1/\gamma_1} \int_0^{x^{-1/\gamma_2}} s^{-2} \alpha_n(\ell_n(s)) ds + O_{\mathbf{p}}(\epsilon) x^{(1-\eta)/\gamma \mp \epsilon}.$$

Therefore

$$\begin{aligned} \frac{\sqrt{k} S_{n1}(x)}{\Lambda(a_k x)} &= \frac{\gamma}{\gamma_1} x^{1/\gamma} \alpha_n\left(\frac{n}{k} \overline{F}(a_k x)\right) \\ &\quad - \frac{\gamma}{\gamma_1} x^{1/\gamma_1} \int_0^{x^{-1/\gamma_2}} s^{-2} \alpha_n(\ell_n(s)) ds + O_{\mathbf{p}}(\epsilon) x^{(1-\eta)/\gamma \mp \epsilon}. \end{aligned}$$

Recall that  $\gamma_1 < \gamma_2$  and  $\gamma/\gamma_2 = \gamma_1/(\gamma_1 + \gamma_2)$ , then we may choose the constant  $\eta$  in such a way that  $\gamma/\gamma_2 < \eta < 1/2$ . Making use of weak approximation (5.12), we obtain

$$\begin{aligned} & \frac{\sqrt{k}S_{n1}(x)}{\Lambda(a_kx)} \\ &= \frac{\gamma}{\gamma_1}x^{1/\gamma}\mathbf{W}\left(\frac{n}{k}\overline{F}(a_kx)\right) + \frac{\gamma}{\gamma_1}x^{1/\gamma}\int_0^{x^{-1/\gamma_2}}s^{-2}\mathbf{W}(\ell_n(s))ds + O_{\mathbf{P}}(\epsilon)x^{(1-\eta)/\gamma\mp\epsilon}. \end{aligned}$$

Note that  $k/n = \overline{F}(G^{\leftarrow}(1 - \overline{G}(a_k)))$ , hence

$$\ell_n(s) = \frac{\overline{F}(G^{\leftarrow}(1 - s\overline{G}(a_k)))}{\overline{F}(G^{\leftarrow}(1 - \overline{G}(a_k)))}.$$

Since  $s \rightarrow \overline{F} \circ G^{\leftarrow}(1 - s)$  is regularly varying at infinity with index  $\gamma_2/\gamma$ , then, from Proposition B.1.10 in [de Haan and Ferreira \(2006\)](#), we have for all large  $n$

$$\omega_n(s) := |\ell_n(s) - s^{\gamma_2/\gamma}| \leq \epsilon s^{\gamma_2/\gamma \pm \epsilon}. \quad (5.17)$$

Recall that  $x_0 > 0$  is fixed, then  $\sup_{x \geq x_0} \sup_{0 < s \leq x^{-1/\gamma_2}} \omega_n(s) \rightarrow 0$ , as  $n \rightarrow \infty$ . On the other hand, by using Levy's modulus of continuity of the Wiener process (see, e.g., Theorem 1.1.1 in [Csörgő and Révész, 1981](#)), we have

$$|\mathbf{W}(\ell_n(s)) - \mathbf{W}(s^{\gamma_2/\gamma})| \leq 2\sqrt{\omega_n(s) \log(1/\omega_n(s))},$$

uniformly on  $s \geq x^{-1/\gamma_2}$ , almost surely. By using the fact that,  $\log u < \epsilon u^{-\epsilon}$  as  $u \downarrow 0$ , together with inequality (5.17), we get  $|\mathbf{W}(\ell_n(s)) - \mathbf{W}(s^{\gamma_2/\gamma})| \leq 2\epsilon s^{(\gamma_2/\gamma)(1-\epsilon)/2}$ . Following our convention, we may write that  $(\gamma_2/\gamma \pm \epsilon)(1 - \epsilon/2) \equiv \gamma_2/\gamma \pm \epsilon$ . Since  $\gamma_1 < \gamma_2$  then  $\gamma_2/(2\gamma) > 1$  and after elementary calculation, we show that uniformly on  $x \geq x_0$

$$\frac{\gamma}{\gamma_1}x^{1/\gamma_1}\int_0^{x^{-1/\gamma_2}}s^{-2}\mathbf{W}(\ell_n(s))ds = \frac{\gamma}{\gamma_1}x^{1/\gamma_1}\int_0^{x^{-1/\gamma_2}}s^{-2}\mathbf{W}(s^{\gamma_2/\gamma})ds + O_{\mathbf{P}}(\epsilon)x^{1/(2\gamma)\pm\epsilon}.$$

By similar arguments, we get

$$\frac{\gamma}{\gamma_1}x^{1/\gamma}\mathbf{W}\left(\frac{n}{k}\overline{F}(a_kx)\right) = \frac{\gamma}{\gamma_1}x^{1/\gamma}\mathbf{W}(x^{-1/\gamma}) + O_{\mathbf{P}}(\epsilon)x^{1/(2\gamma)\pm\epsilon}.$$

It is obvious that  $O_{\mathbf{P}}(\epsilon)x^{1/(2\gamma)\pm\epsilon} + O_{\mathbf{P}}(\epsilon)x^{(1-\eta)/\gamma\pm\epsilon} = O_{\mathbf{P}}(\epsilon)x^{(1-\eta)/\gamma\pm\epsilon}$ , it follows that

$$\frac{\sqrt{k}S_{n1}(x)}{\Lambda(a_kx)} = \frac{\gamma}{\gamma_1}x^{1/\gamma}\mathbf{W}(x^{-1/\gamma}) + \frac{\gamma}{\gamma_1}x^{1/\gamma_1}\int_0^{x^{-1/\gamma_2}}s^{-2}\mathbf{W}(s^{\gamma_2/\gamma})ds + O_{\mathbf{P}}(\epsilon)x^{(1-\eta)/\gamma\pm\epsilon}.$$

After a change of variables, this may be rewritten into

$$\begin{aligned} & \frac{\sqrt{k}S_{n1}(x)}{\Lambda(a_kx)} \\ &= \frac{\gamma}{\gamma_1}x^{1/\gamma}\mathbf{W}(x^{-1/\gamma}) + \frac{\gamma}{\gamma_1 + \gamma_2}x^{1/\gamma} \int_0^1 t^{-\gamma/\gamma_2-1}\mathbf{W}(x^{-1/\gamma}t) dt + O_{\mathbf{P}}(\epsilon)x^{(1-\eta)/\gamma \pm \epsilon}. \end{aligned} \quad (5.18)$$

Now, we consider the second term  $S_{n2}(x)$ . We have  $\bar{F}_n(z) = 0$ , for  $z \geq X_{n:n}$ , thus

$$S_{n2}(x) = \int_{xX_{n-k:n}}^{X_{n:n}} \frac{C_n(z) - C(z)}{C_n(z)C(z)} d\bar{F}_n(z).$$

Therefore

$$|S_{n2}(x)| \leq \theta_n \int_{xX_{n-k:n}}^{\infty} \frac{|C_n(z) - C(z)|}{C^2(z)} dF_n(z),$$

where  $\theta_n := \sup_{X_{1:n} \leq z \leq X_{n:n}} \{C(z)/C_n(z)\}$ , which is stochastically bounded (see, e.g., [Stute and Wang, 2008](#)). By recalling that  $C = \bar{G} - \bar{F}$  and  $C_n = \bar{G}_n - \bar{F}_n$ , with  $G_n$  denoting the empirical df of  $G$ , we write  $|S_{n2}(x)| \leq \theta_n (T_{n1}(x) + T_{n2}(x))$ , where

$$T_{n1}(x) := \int_{xX_{n-k:n}}^{\infty} \frac{|\bar{F}_n(z) - \bar{F}(z)|}{C^2(z)} dF_n(z)$$

and

$$T_{n2}(x) := \int_{xX_{n-k:n}}^{\infty} \frac{|\bar{G}_n(z) - \bar{G}(z)|}{C^2(z)} dF_n(z).$$

It is easy to verify that, by a change of variables, we have

$$\begin{aligned} \frac{\sqrt{k}T_{n1}(x)}{\Lambda(a_kx)} &= d_n(x) \frac{k/n}{C(a_k)} \frac{C(a_kx)}{C(a_k)} \\ &\times \frac{C^2(a_k)}{C^2(xX_{n-k:n})} \int_1^{\infty} \frac{\left| \alpha_n \left( \frac{n}{k} \bar{F}(xX_{n-k:n}z) \right) \right|}{C^2(xX_{n-k:n}z)/C^2(xX_{n-k:n})} d \frac{F_n(xX_{n-k:n}z)}{\bar{F}(a_k)}. \end{aligned}$$

Recall that, uniformly on  $x \geq x_0$ , we have  $C(a_k)/C(xX_{n-k:n}) = O_{\mathbf{P}}(1)x^{1/\gamma_2 \pm \epsilon}$ .

Moreover, we use (5.16) and (5.15) to write

$$\begin{aligned} \frac{\sqrt{k}T_{n1}(x)}{\Lambda(a_kx)} &= O_{\mathbf{P}}\left(\frac{k/n}{C(a_k)}\right) x^{1/\gamma \pm \epsilon} \\ &\times \int_1^{\infty} z^{2/\gamma_2} \left| \alpha_n \left( \frac{n}{k} \bar{F}(xX_{n-k:n}z) \right) \right| d \frac{F_n(xX_{n-k:n}z)}{\bar{F}(a_k)}. \end{aligned}$$

On the other hand, by using the stochastic boundedness of  $\sup_{0 < s \leq 1} |\alpha_n(s)|/s^\eta$  we get

$$\frac{\sqrt{k}T_{n1}(x)}{\Lambda(a_kx)} = O_{\mathbf{P}}\left(\frac{k/n}{C(a_k)}\right) x^{(1-\eta)/\gamma \pm \epsilon} \int_1^{\infty} z^{2/\gamma_2 - \eta/\gamma \pm \epsilon} d \frac{\bar{F}_n(xX_{n-k:n}z)}{\bar{F}(a_k)},$$

where the integral may be split as follows

$$\int_1^\infty z^{2/\gamma_2 - \eta/\gamma \pm \epsilon} d \frac{\overline{F}_n(xX_{n-k:n}z)}{\overline{F}(a_k)} = P_n(x) + Q_n(x),$$

where

$$P_n(x) := \int_1^\infty z^{2/\gamma_2 - \eta/\gamma \pm \epsilon} d \left\{ \frac{\overline{F}_n(xX_{n-k:n}z) - \overline{F}(xX_{n-k:n}z)}{\overline{F}(a_k)} \right\},$$

and

$$Q_n(x) := \int_1^\infty z^{2/\gamma_2 - \eta/\gamma \pm \epsilon} d \frac{\overline{F}(xX_{n-k:n}z)}{\overline{F}(a_k)}.$$

It is clear that

$$P_n(x) = k^{-1/2} \int_1^\infty z^{2/\gamma_2 - \eta/\gamma \pm \epsilon} d\alpha_n \left( \frac{n}{k} \overline{F}(xX_{n-k:n}z) \right).$$

By similar arguments as those used above, we show that  $P_n(x) = o_{\mathbf{P}}(x^{-\eta/\gamma \pm \epsilon})$  and  $Q_n(x) = O_{\mathbf{P}}(x^{-1/\gamma \pm \epsilon})$ . Therefore

$$\frac{\sqrt{k}T_{n1}(x)}{\Lambda(a_k x)} = x^{-\eta/\gamma \pm \epsilon} O_{\mathbf{P}} \left( \frac{k/n}{C(a_k)} \right).$$

Next, let  $V_i := \overline{G}(Y_i)$ ,  $i = 1, \dots, n$ , and define the corresponding tail empirical process  $\beta_n(s) := \sqrt{k}(\mathbf{V}_n(s) - s)$ , for  $0 \leq s \leq 1$ , where  $\mathbf{V}_n(s) := k^{-1} \sum_{i=1}^n \mathbf{1}(V_i < ks/n)$ . Like for  $\alpha_n(s)$ , we also have  $\sup_{0 < s \leq 1} |\beta_n(s)|/s^\eta = O_{\mathbf{P}}(1)$ , therefore by similar arguments as those used for  $T_{n1}(x)$ , with the facts that  $\overline{G}(t) \sim C(t)$  as  $t \rightarrow \infty$  and  $\gamma_2 > \gamma$ , we show that

$$\frac{\sqrt{k}T_{n2}(x)}{\Lambda(a_k x)} = O_{\mathbf{P}} \left( \frac{k/n}{C^{1-\eta}(a_k)} \right) x^{(1-\eta)/\gamma \pm \epsilon}.$$

From Lemma 6.2 (ii), we have that both  $\frac{n}{k}C(a_k)$  and  $\frac{n}{k}C^{1-\eta}(a_k)$  tend to infinity, it follows that

$$\frac{\sqrt{k}T_{n1}(x)}{\Lambda(a_k x)} = o_{\mathbf{P}}(x^{-\eta/\gamma \pm \epsilon}) \quad \text{and} \quad \frac{\sqrt{k}T_{n2}(x)}{\Lambda(a_k x)} = o_{\mathbf{P}}(x^{(1-\eta)/\gamma \pm \epsilon}).$$

Since  $o_{\mathbf{P}}(x^{-\eta/\gamma \pm \epsilon}) + o_{\mathbf{P}}(x^{(1-\eta)/\gamma \pm \epsilon}) = o_{\mathbf{P}}(x^{(1-\eta)/\gamma \pm \epsilon})$ , then

$$\frac{\sqrt{k}S_{n2}(x)}{\Lambda(a_k x)} = o_{\mathbf{P}}(x^{(1-\eta)/\gamma \pm \epsilon}). \quad (5.19)$$

Let us now focus on the third term  $S_{n3}$ , which, by integration by parts, equals the sum of

$$S_{n3}^{(1)}(x) := - \int_{xX_{n-k:n}}^{xa_k} \frac{\overline{F}_n(z) - \overline{F}(z)}{C^2(z)} dC(z),$$

and

$$S_{n3}^{(2)}(x) = - \frac{\overline{F}_n(a_k x) - \overline{F}(a_k x)}{C(a_k x)} + \frac{\overline{F}_n(xX_{n-k:n}) - \overline{F}(xX_{n-k:n})}{C(xX_{n-k:n})}.$$



By using the change of variables  $z = txa_k$  we get

$$\frac{\sqrt{k}S_{n3}^{(1)}(x)}{\Lambda(a_kx)} = -d_n(x) \int_{X_{n-k:n}/a_k}^1 \frac{\alpha_n\left(\frac{n}{k}\bar{F}(a_kxz)\right)}{(C(a_kxz)/C(a_kx))^2} d\frac{C(a_kxz)}{C(a_kx)},$$

and

$$\frac{\sqrt{k}S_{n3}^{(2)}(x)}{\Lambda(a_kx)} = -d_n(x) \left\{ \alpha_n\left(\frac{n}{k}\bar{F}(a_kx)\right) - \frac{C(a_kx)}{C(xX_{n-k:n})} \alpha_n\left(\frac{n}{k}\bar{F}(xX_{n-k:n})\right) \right\}.$$

Routine manipulations, including Proposition B.1.10 in [de Haan and Ferreira \(2006\)](#)

and the stochastic boundedness of  $\sup_{0 < s \leq 1} |\alpha_n(s)|/s^\eta$ , yield

$$\frac{\sqrt{k}S_{n3}^{(1)}(x)}{\Lambda(a_kx)} = o_{\mathbf{P}}(x^{(1-\eta)/\gamma \pm \epsilon}) \quad \text{and} \quad \frac{\sqrt{k}S_{n3}^{(2)}(x)}{\Lambda(a_kx)} = o_{\mathbf{P}}(x^{(1-\eta)/\gamma \pm \epsilon}).$$

It follows that

$$\sqrt{k}S_{n3}(x)/\Lambda(a_kx) = o_{\mathbf{P}}(x^{(1-\eta)/\gamma \pm \epsilon}). \quad (5.20)$$

By gathering results (5.18), (5.19) and (5.20), we obtain

$$\begin{aligned} & \sqrt{k} \frac{\Lambda_n(xX_{n-k:n}) - \Lambda(xX_{n-k:n})}{\Lambda(a_kx)} \\ &= \frac{\gamma}{\gamma_1} x^{1/\gamma} \mathbf{W}(x^{-1/\gamma}) + \frac{\gamma}{\gamma_1 + \gamma_2} x^{1/\gamma} \int_0^1 t^{-\gamma/\gamma_2 - 1} \mathbf{W}(x^{-1/\gamma}t) dt + O_{\mathbf{P}}(\epsilon) x^{(1-\eta)/\gamma \pm \epsilon}, \end{aligned} \quad (5.21)$$

which yields that

$$\begin{aligned} & x^{1/\gamma_1} \sqrt{k} \mathbf{M}_{n1}^*(x) \\ &= x^{1/\gamma} \left\{ \frac{\gamma}{\gamma_1} \mathbf{W}(x^{-1/\gamma}) + \frac{\gamma}{\gamma_1 + \gamma_2} \int_0^1 t^{-\gamma/\gamma_2 - 1} \mathbf{W}(x^{-1/\gamma}t) dt \right\} + O_{\mathbf{P}}(\epsilon) x^{(1-\eta)/\gamma \pm \epsilon}. \end{aligned}$$

We show that the expectation of the absolute value of the first term in the right-hand side of the previous equation equals  $O_{\mathbf{P}}(x^{1/(2\gamma)})$ . Since  $1/(2\gamma) < (1-\eta)/\gamma$ , we have  $x^{1/\gamma_1} \sqrt{k} \mathbf{M}_{n1}^*(x) = O_{\mathbf{P}}(x^{(1-\eta)/\gamma \pm \epsilon})$ , which leads to

$$x^{1/\gamma_1} \sqrt{k} \mathbf{M}_{n1}(x) = x^{1/\gamma_1} \sqrt{k} \mathbf{M}_{n1}^*(x) + o_{\mathbf{P}}(x^{(1-\eta)/\gamma \pm \epsilon}).$$

Recall that  $\epsilon > 0$  is chosen sufficiently small, then for any  $0 < \eta < 1/2$ , we have

$$\begin{aligned} & x^{1/\gamma_1} \sqrt{k} \mathbf{M}_{n1}(x) \\ &= x^{1/\gamma} \left\{ \frac{\gamma}{\gamma_1} \mathbf{W}(x^{-1/\gamma}) + \frac{\gamma}{\gamma_1 + \gamma_2} \int_0^1 t^{-\gamma/\gamma_2 - 1} \mathbf{W}(x^{-1/\gamma}t) dt \right\} + O_{\mathbf{P}}(\epsilon) x^{(1-\eta)/\gamma \pm \epsilon}. \end{aligned}$$

Before we treat the term  $\mathbf{M}_{n2}(x)$ , it is worth mentioning that by letting  $x = 1$  in the previous approximation, we infer that

$$\frac{\bar{\mathbf{F}}_n(X_{n-k:n})}{\bar{\mathbf{F}}(X_{n-k:n})} - 1 = O_{\mathbf{P}}(k^{-1/2}) = o_{\mathbf{P}}(1), \quad (k \rightarrow \infty). \quad (5.22)$$

This, with the regular variation of  $\overline{\mathbf{F}}$ , imply that

$$\frac{\overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}_n(X_{n-k:n})} = (1 + O_{\mathbf{P}}(x^{\pm\epsilon})) x^{-1/\gamma_1}. \quad (5.23)$$

To represent  $\sqrt{k}\mathbf{M}_{n2}(x)$ , we apply results (5.21) (for  $x = 1$ ) and (5.23) to get

$$x^{1/\gamma_1}\sqrt{k}\mathbf{M}_{n2}(x) = - \left\{ \frac{\gamma}{\gamma_1} \mathbf{W}(1) + \frac{\gamma}{\gamma_1 + \gamma_2} \int_0^1 t^{-\gamma/\gamma_2 - 1} \mathbf{W}(t) dt \right\} + O_{\mathbf{P}}(\epsilon) x^{\pm\epsilon}.$$

For the third term  $\mathbf{M}_{n3}(x)$ , we write

$$x^{1/\gamma_1}\sqrt{k}\mathbf{M}_{n3}(x) = \left( \frac{\overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}_n(X_{n-k:n})} - x^{-1/\gamma_1} \right) x^{1/\gamma_1}\sqrt{k}\mathbf{M}_{n1}(x),$$

which, by equation (5.23), is equal to  $O_{\mathbf{P}}(\epsilon) x^{-1/\gamma_1 + (1-\eta)/\gamma \pm \epsilon}$ . Let  $\eta_0$  be such that  $\gamma/\gamma_2 < \eta_0 < \eta < 1/2$ , then  $\eta_0 - \eta < 0$  and for  $\epsilon > 0$  sufficiently small, we have  $(\eta_0 - \eta)/\gamma + \epsilon < 0$ . Since  $x \geq x_0 > 0$ , then  $O_{\mathbf{P}}(\epsilon) x^{(\eta_0 - \eta)/\gamma \pm \epsilon} = O_{\mathbf{P}}(\epsilon)$  and thus

$$x^{1/\gamma_1 - (1-\eta_0)/\gamma} \left\{ \sqrt{k}(\mathbf{M}_{n1}(x) + \mathbf{M}_{n2}(x) + \mathbf{M}_{n3}(x)) - \mathbf{\Gamma}(x; \mathbf{W}) \right\} = O_{\mathbf{P}}(\epsilon), \quad (5.24)$$

where  $\mathbf{\Gamma}(x; \mathbf{W})$  is the Gaussian process given in Theorem 2.1. For the fourth term  $\mathbf{M}_{n4}(x)$ , it suffices to use the uniform inequality to second-order condition (2.9), given in assertion (2.3.23) of Theorem 2.3.9 in de Haan and Ferreira (2006), to get

$$\sqrt{k}\mathbf{M}_{n4}(x) = (1 + o_{\mathbf{P}}(1)) x^{-1/\gamma_1} \frac{x^{\tau_1/\gamma_1} - 1}{\gamma_1 \tau_1} \sqrt{k} \tilde{\mathbf{A}}_{\mathbf{F}}(1/\overline{\mathbf{F}}(X_{n-k:n})),$$

for a possibly different function  $\tilde{\mathbf{A}}_{\mathbf{F}}$  with  $\tilde{\mathbf{A}}_{\mathbf{F}} \sim \mathbf{A}_{\mathbf{F}}$ . Then Proposition B.1.10 in de Haan and Ferreira (2006) and the fact that  $t \rightarrow \tilde{\mathbf{A}}_{\mathbf{F}}(1/\overline{\mathbf{F}}(\mathbb{U}_F(t))) =: \mathbf{A}_0(t)$  is regularly varying with index  $\tau_1/\gamma_1$  with  $X_{n-k:n}/a_k \xrightarrow{\mathbf{P}} 1$ , imply that

$$\frac{\mathbf{A}_0(1/\overline{\mathbf{F}}(X_{n-k:n}))}{\mathbf{A}_0(1/\overline{\mathbf{F}}(a_k))} \xrightarrow{\mathbf{P}} 1, \text{ as } n \rightarrow \infty,$$

as well. Since  $o_{\mathbf{P}}(1) x^{-1/\gamma_1} \frac{x^{\tau_1/\gamma_1} - 1}{\gamma_1 \tau_1} = o_{\mathbf{P}}(x^{-1/\gamma_1 + (1-\eta)/\gamma \pm \epsilon})$ , and by assumption  $\sqrt{k}\mathbf{A}_0(1/\overline{\mathbf{F}}(a_k)) = \sqrt{k}\mathbf{A}_0(n/k) = O(1)$ , it follows that

$$\sqrt{k}\mathbf{M}_{n4}(x) = x^{-1/\gamma_1} \frac{x^{\tau_1/\gamma_1} - 1}{\gamma_1 \tau_1} \sqrt{k}\mathbf{A}_0(n/k) + o_{\mathbf{P}}(x^{-1/\gamma_1 + (1-\eta)/\gamma \pm \epsilon}).$$

Finally, by letting  $\epsilon \downarrow 0$  in (5.24), we end up with

$$\sup_{x \geq x_0} x^{1/\gamma_1 - (1-\eta_0)/\gamma} \left| \mathbf{D}_n(x) - \mathbf{\Gamma}(x; \mathbf{W}) - x^{-1/\gamma_1} \frac{x^{\tau_1/\gamma_1} - 1}{\gamma_1 \tau_1} \sqrt{k}\mathbf{A}_0(n/k) \right| \xrightarrow{\mathbf{P}} 0,$$

for every  $x_0 > 0$  and  $\gamma/\gamma_2 < \eta_0 < \eta < 1/2$ . Letting  $\eta_0 := 1/2 - \xi$  and recalling that  $1/\gamma_1 = 1/\gamma - 1/\gamma_2$  yields that  $0 < \xi < 1/2 - \gamma/\gamma_2$  and achieves the proof.  $\square$

**5.2. Proof of Theorem 3.1.** We start by proving the consistency of  $\hat{\gamma}_1$  that we write as  $\hat{\gamma}_1 = \int_1^\infty x^{-1} \bar{\mathbf{F}}_n(xX_{n-k:n}) / \bar{\mathbf{F}}_n(X_{n-k:n}) dx$ . It is readily checked that this may be decomposed into the sum of

$$I_{1n} := \int_1^\infty x^{-1} \frac{\bar{\mathbf{F}}(xX_{n-k:n})}{\bar{\mathbf{F}}(X_{n-k:n})} dx \text{ and } I_{2n} := \int_1^\infty x^{-1} \sum_{i=1}^3 \mathbf{M}_{ni}(x) dx.$$

By the regular variation of  $\bar{\mathbf{F}}$  (1.1) and Potter's inequalities, we get  $I_{1n} \xrightarrow{\mathbf{P}} \gamma_1$  as  $n \rightarrow \infty$ . Then, we just need to show that  $I_{2n}$  tends to zero in probability. From (5.24) we have

$$I_{2n} = \frac{1}{\sqrt{k}} \int_1^\infty x^{-1} \mathbf{\Gamma}(x; \mathbf{W}) dx + \frac{1}{\sqrt{k}} \int_1^\infty x^{-1} o_{\mathbf{P}}(x^{(1-\eta)/\gamma-1/\gamma_1}) dx.$$

On the one hand, since  $\gamma/\gamma_2 < \eta$ , the second integral above is finite and therefore the second term of  $I_{2n}$  is negligible in probability. On the other hand, we have

$$\begin{aligned} \int_1^\infty x^{-1} \mathbf{\Gamma}(x; \mathbf{W}) dx &= \frac{\gamma}{\gamma_1} \int_1^\infty x^{1/\gamma_2-1} \{ \mathbf{W}(x^{-1/\gamma}) - x^{-1/\gamma} \mathbf{W}(1) \} dx + \frac{\gamma}{\gamma_1 + \gamma_2} \\ &\quad \times \int_1^\infty x^{1/\gamma_2-1} \left\{ \int_0^1 s^{-\gamma/\gamma_2-1} \{ \mathbf{W}(x^{-1/\gamma}s) - x^{-1/\gamma} \mathbf{W}(s) \} ds \right\} dx, \end{aligned}$$

which, after some elementary but tedious manipulations of integral calculus (change of variables and integration by parts), becomes

$$\begin{aligned} \int_1^\infty x^{-1} \mathbf{\Gamma}(x; \mathbf{W}) dx &= -\gamma \mathbf{W}(1) \\ &\quad + \frac{\gamma}{\gamma_1 + \gamma_2} \int_0^1 (\gamma_2 - \gamma_1 - \gamma \log s) s^{-\gamma/\gamma_2-1} \mathbf{W}(s) ds. \end{aligned} \tag{5.25}$$

By using the facts that  $\mathbf{E}|\mathbf{W}(s)| \leq s^{1/2}$  and  $\gamma_1 < \gamma_2$ , we deduce that  $\int_1^\infty x^{-1} \mathbf{\Gamma}(x; \mathbf{W}) dx$  is stochastically bounded and therefore the first term of  $I_{2n}$  is negligible in probability as well. Consequently, we have  $I_{2n} = o_{\mathbf{P}}(1)$  when  $n \rightarrow \infty$ , as sought. As for the Gaussian representation result, it is easy to verify that  $\sqrt{k}(\hat{\gamma}_1 - \gamma_1) = \int_1^\infty x^{-1} \mathbf{D}_n(x) dx$ . Then, applying Theorem 2.1 yields that

$$\sqrt{k}(\hat{\gamma}_1 - \gamma_1) = \frac{\sqrt{k} \mathbf{A}_0(n/k)}{1 - \tau} + \int_1^\infty x^{-1} \mathbf{\Gamma}(x; \mathbf{W}) dx + o_{\mathbf{P}}(1),$$

and finally, using result (5.25) completes the proof.  $\square$

**5.3. Proof of Corollary 3.1.** We set

$$\sqrt{k}(\hat{\gamma}_1 - \gamma_1) = \gamma \Delta + \frac{\sqrt{k} \mathbf{A}_0(n/k)}{1 - \tau} + o_{\mathbf{P}}(1),$$

where  $\Delta := a\Delta_1 + b\Delta_2 - \Delta_3$ , with  $a := (\gamma_2 - \gamma_1) / (\gamma_1 + \gamma_2)$ ,  $b := -\gamma / (\gamma_1 + \gamma_2)$  and

$$\Delta_1 := \int_0^1 s^{\rho-2} \mathbf{W}(s) ds, \quad \Delta_2 := \int_0^1 s^{\rho-2} \mathbf{W}(s) \log s ds, \quad \Delta_3 := \mathbf{W}(1),$$

with  $\rho := 1 - \gamma/\gamma_2 > 0$ .

It is clear that the asymptotic mean is equal to  $\lim_{n \rightarrow \infty} \sqrt{k} \mathbf{A}_0(n/k) / (1 - \tau)$ , while for the asymptotic variance we find, after elementary but tedious computations, the following covariances:

$$\begin{aligned} \mathbf{E}[\Delta_1^2] &= \frac{2}{\rho(2\rho-1)}, \quad \mathbf{E}[\Delta_2^2] = \frac{2(4\rho-1)}{\rho^2(2\rho-1)^3}, \quad \mathbf{E}[\Delta_3^2] = 1, \\ \mathbf{E}[\Delta_1\Delta_2] &= \frac{1-4\rho}{\rho^2(2\rho-1)^2}, \quad \mathbf{E}[\Delta_1\Delta_3] = \frac{1}{\rho}, \quad \mathbf{E}[\Delta_2\Delta_3] = -\frac{1}{\rho^2}. \end{aligned}$$

It follows that

$$\mathbf{E}[\Delta^2] = \frac{2a^2}{\rho(2\rho-1)} + \frac{2b^2(4\rho-1)}{\rho^2(2\rho-1)^3} + \frac{2ab(1-4\rho)}{\rho^2(2\rho-1)^2} + \frac{2b}{\rho^2} - \frac{2a}{\rho} + 1.$$

Replacing  $a$ ,  $b$  and  $\rho$  by their values achieves the proof.  $\square$

## Concluding notes

We would like to emphasize the fact that, unlike [Gardes and Stupfler \(2015\)](#) who defined their estimator in terms of two (not necessarily equal) sample fractions  $k = k'$  of upper order statistics from  $X$  and  $Y$  respectively, we build our estimator on the basis of just a single sample fraction. The consideration of two distinct sample fractions poses a problem from a computational point of view, as the issue of selecting an optimal couple of sample fractions is not as easy and usual as determining just one best number of top statistics to be used in parameter estimate computation. Besides that, [Gardes and Stupfler \(2015\)](#) didn't treat the asymptotic normality when  $k = k'$  and only carried out their simulations in this very particular case, as they mentioned in their conclusion. For these reasons, we don't compare the two estimators in [Section 5](#).

A more thorough simulation study, with confidence interval construction and eventual comparison with the estimator of [Gardes and Stupfler \(2015\)](#), will be part of a future work. Another point, beyond the scope of the present paper, that deserves to be considered is to reduce estimation biases under random truncation. Similar anterior works were done with complete datasets by, for instance, [Peng and Qi \(2004\)](#), [Li and Peng \(2010\)](#) and [Brahimi et al. \(2013\)](#).

We finish this work by making a comment on relation (3.11), which actually is a special case of a more general functional of the distribution tail defined by

$$\Gamma_t(g, \alpha) := \frac{\frac{1}{\overline{F}(t)} \int_t^\infty g\left(\frac{\overline{F}(x)}{\overline{F}(t-)}\right) \left(\log \frac{x}{t}\right)^\alpha dF(x)}{\int_0^1 g(x) (-\log x)^\alpha dx}, \quad t \geq 0,$$

where  $g$  is some weight function and  $\alpha$  some positive real number. As a consequence of the fact that  $\lim_{t \rightarrow \infty} \Gamma_t(g, \alpha) = \gamma^\alpha$ , this functional can be considered as the starting point to constructing a whole class of estimators for distribution tail parameters. Indeed, in the complete data case, we replace  $F$  by its empirical counterpart  $F_n$  and  $t$  by  $X_{n-k:n}$  to get the following statistic which generalizes several extreme value theory based procedures of estimation already existing in the literature:

$$\Gamma_{n,k}(g, \alpha) := \frac{\frac{1}{k} \sum_{i=1}^k g\left(\frac{i}{k+1}\right) \left(\log \frac{X_{n-i+1:n}}{X_{n-k:n}}\right)^\alpha}{\int_0^1 g(x) (-\log x)^\alpha dx}.$$

When  $g = \alpha = 1$ , we recover the famous Hill estimator (Hill, 1975). For a detailed list of extreme value index estimators drawn from the statistic above, we refer to the paper of Ciuperca and Mercadier (2010), where the authors propose an estimation approach of the second-order parameter by considering differences and quotients of several forms of  $\Gamma_{n,k}(g, \alpha)$ . By analogy, when we deal with randomly truncated observations, we substitute the product-limit estimator  $\mathbf{F}_n$  for  $F$  in the formula of  $\Gamma_t(g, \alpha)$  in order to obtain the following family of parameter estimators under random truncation:

$$\Gamma_{n,k}(g, \alpha) := \frac{\sum_{i=1}^k a_n^{(i)} g\left(\frac{\overline{\mathbf{F}}_n(X_{n-i+1:n})}{\overline{\mathbf{F}}_n(X_{n-k+1:n})}\right) \left(\log \frac{X_{n-i+1:n}}{X_{n-k:n}}\right)^\alpha}{\sum_{i=1}^k a_n^{(i)} \int_0^1 g(x) (-\log x)^\alpha dx}.$$

This would have fruitful consequences on the statistical analysis of extremes under random truncation.

## REFERENCES

- Brahimi, B., Meraghni, D., Necir, A., Yahia, D., 2013. A bias-reduced estimator for the mean of a heavy-tailed distribution with an infinite second moment. *J. Statist. Plann. Inference* 143, no. 6, 1064-1081.
- Ciuperca, G., Mercadier, C., 2010. Semi-parametric estimation for heavy tailed distributions. *Extremes* 13, 55-87.

- Csörgő, M., Révész, P., 1981. Strong approximations in probability and statistics. Probability and Mathematical Statistics. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London.
- Dree, H., de Haan, L., Li, D., 2006. Approximations to the tail empirical distribution function with application to testing extreme value conditions. *J. Statist. Plann. Inference* 136, 3498-3538.
- Einmahl, J.H.J., 1992. Limit theorems for tail processes with application to intermediate quantile estimation. *J. Statist. Plann. Inference* 32, 137-145.
- Einmahl, J.H.J., de Haan, L., Li, D., 2006. Weighted approximations of tail copula processes with application to testing the bivariate extreme value condition. *Ann. Statist.* 34, 1987-2014.
- Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. Modelling Extremal Events for Insurance and Finance. Springer, Berlin.
- Escudero, F., Ortega, E., 2008. Actuarial comparisons for aggregate claims with randomly right-truncated claims. *Insurance Math. Econom.* 43, 255-262.
- Gardes, L., Stupfler, G., 2015. Estimating extreme quantiles under random truncation. *TEST* 24, 207-227.
- de Haan, L., Stadtmüller, U., 1996. Generalized regular variation of second order. *J. Australian Math. Soc. (Series A)* 61, 381-395.
- de Haan, L., Ferreira, A., 2006. Extreme Value Theory: An Introduction. *Springer*.
- Herbst, T., 1999. An application of randomly truncated data models in reserving IBNR claims. *Insurance Math. Econom.* 25, 123-131.
- Hill, B.M., 1975. A simple general approach to inference about the tail of a distribution. *Ann. Statist.* 3, 1163-1174.
- Hua, L., Harry, J., 2011. Second order regular variation and conditional tail expectation of multiple risks. *Insurance Math. Econom.* 49: 537-546.
- Koning, A.J., Peng, L., 2008. Goodness-of-fit tests for a heavy tailed distribution. *J. Statist. Plann. Inference* 138, no. 12, 3960-3981.
- Lawless, J.F., 2002. Statistical Models and Methods for Lifetime Data, Second Edition. Wiley Series in Probability and Statistics.
- Li, D., Peng, L., Yang, J., 2010. Bias reduction for high quantiles. *J. Statist. Plann. Inference* 140, no. 9, 2433-2441.
- Lynden-Bell, D., 1971. A method of allowing for known observational selection in small samples applied to 3CR quasars. *Monthly Notices Roy. Astronom. Soc.* 155: 95-118.

- Mason, D.M., 1982. Laws of large numbers for sums of extreme values. *Ann. Probab.* 10, 756-764.
- Peng, L., Qi, Y., 2004. Estimating the first- and second-order parameters of a heavy-tailed distribution. *Aust. N. Z. J. Stat.* 46, no. 2, 305-312.
- Reiss, R.D., Thomas, M., 2007. *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*, 3rd ed. Birkhäuser Verlag, Basel, Boston, Berlin.
- Resnick, S., 2006. *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer.
- Strzalkowska-Kominiak, E., Stute, W., 2009. Martingale representations of the Lynden-Bell estimator with applications. *Statist. Probab. Lett.* 79: 814-820.
- Stute, W., Wang, J., 2008. The central limit theorem under random truncation. *Bernoulli* 14, 604-622.
- Wang M.C., Jewell N.P., Tsai, W.Y., 1986. Asymptotic properties of the product-limit estimate under random truncation. *Ann Statist* 14:1597-1605.
- Wang, M.C., 1989. A semiparametric model for randomly truncated data. *Statist. Assoc.*, 84, 742-748
- Woodroffe, M., 1985. Estimating a distribution function with truncated data. *Ann. Statist.* 13:163-177.

## 6. Appendix

**Lemma 6.1.** *Assume that both second-order conditions (2.9) and (2.10) hold. Then, for all large  $x$ , there exist constants  $d_1, d_2 > 0$ , such that*

$$\overline{F}(x) = (1 + o(1)) d_1 x^{-1/\gamma} \text{ and } \overline{G}(x) = (1 + o(1)) d_2 x^{-1/\gamma_2}.$$

*Proof.* We only show the first statement since the second one follows by similar arguments. To this end, we rewrite the first equation of (1.2) into

$$\overline{F}(x) = -p^{-1} \overline{G}(x) \overline{F}(x) \int_1^\infty \frac{\overline{G}(xz)}{\overline{G}(x)} d \frac{\overline{F}(xz)}{\overline{F}(x)}.$$

By applying Proposition B.1.10 in de Haan and Ferreira (2006) to both  $\overline{F}$  and  $\overline{G}$ , it is easy to check that

$$\int_1^\infty \frac{\overline{G}(xz)}{\overline{G}(x)} d \frac{\overline{F}(xz)}{\overline{F}(x)} = -(1 + o(1)) \gamma / \gamma_1.$$

On the other hand, since  $\overline{F}$  and  $\overline{G}$  satisfy the aforementioned second-order conditions, then in view of Lemma 3 in Hua and Joe (2011), there exist two constants  $a_1, a_2 > 0$ , such that  $\overline{F}(x) = (1 + o(1)) a_1 x^{-1/\gamma_1}$  and  $\overline{G}(x) = (1 + o(1)) a_2 x^{-1/\gamma_2}$ , as  $x \rightarrow \infty$ . Therefore  $\overline{F}(x) = (1 + o(1)) d_1 x^{-1/\gamma}$  with  $d_1 = p^{-1} a_1 a_2 \gamma / \gamma_1$ .  $\square$

**Lemma 6.2.** *Under the assumptions of Lemma 6.1, we have*

- (i)  $\lim_{t \rightarrow \infty} C(t) / \overline{G}(t) = 1$ .
- (ii)  $\lim_{t \rightarrow \infty} t^{1/\nu} C(\mathbb{U}_F(t)) = \infty$ , for each  $0 < \nu \leq 1$ .
- (iii)  $\lim_{t \rightarrow \infty} \sup_{x \geq x_0} x^{-1/\gamma \pm \epsilon} \left| (t \Lambda(x \mathbb{U}_F(t)) C(x \mathbb{U}_F(t)))^{-1} - (\gamma / \gamma_1) x^{1/\gamma} \right| = 0$ ,  
for  $x_0 > 0$  and any sufficiently small  $\epsilon > 0$ .

*Proof.* For assertion (i), write  $C(t) / \overline{G}(t) = 1 - \overline{F}(t) / \overline{G}(t)$  and observe that from Lemma 6.1 we have  $\overline{F}(t) / \overline{G}(t) = (1 + o(1)) (d_1 / d_2) t^{1/\gamma_2 - 1/\gamma}$ . Since  $1/\gamma_2 - 1/\gamma < 0$ , then  $\overline{F}(t) / \overline{G}(t) = o(1)$ , that is  $C(t) / \overline{G}(t) = 1 + o(1)$  as sought. For result (ii), Lemma 6.1 implies that  $\mathbb{U}_F(t) = (1 + o(1)) (d_1 t)^\gamma$  (as  $t \rightarrow \infty$ ), it follows that  $C(\mathbb{U}_F(t)) = (1 + o(1)) d_2 (d_1 t)^{-\gamma/\gamma_2}$ . Since  $0 < \gamma/\gamma_2 < 1$ , then for every  $0 < \nu \leq 1$ ,  $t^{1/\nu} C(\mathbb{U}_F(t)) \rightarrow \infty$  as  $t \rightarrow \infty$ . To prove (iii), we first show that

$$t \Lambda(x \mathbb{U}_F(t)) C(x \mathbb{U}_F(t)) - (\gamma_1 / \gamma) x^{-1/\gamma} = O(\epsilon) x^{-1/\gamma \pm \epsilon}. \quad (6.26)$$

Recalling that  $\Lambda(x) = \int_x^\infty dF(z) / C(z)$  and  $\overline{F}(\mathbb{U}_F(t)) = 1/t$ , we write

$$t \Lambda(x \mathbb{U}_F(t)) C(x \mathbb{U}_F(t)) = - \frac{C(x \mathbb{U}_F(t))}{C(\mathbb{U}_F(t))} \int_x^\infty \frac{C(\mathbb{U}_F(t))}{C(z \mathbb{U}_F(t))} \frac{d\overline{F}(z \mathbb{U}_F(t))}{\overline{F}(\mathbb{U}_F(t))}. \quad (6.27)$$



Observe now that  $t\Lambda(x\mathbb{U}_F(t))C(x\mathbb{U}_F(t)) - \frac{\gamma_1}{\gamma}x^{-1/\gamma}$  may be decomposed into the sum of

$$D_1(s; t) := - \left( \frac{C(x\mathbb{U}_F(t))}{C(\mathbb{U}_F(t))} - x^{-1/\gamma_2} \right) \int_x^\infty \frac{C(\mathbb{U}_F(t))}{C(z\mathbb{U}_F(t))} \frac{d\overline{F}(z\mathbb{U}_F(t))}{\overline{F}(\mathbb{U}_F(t))},$$

$$D_2(s; t) := -x^{-1/\gamma_2} \int_x^\infty \left( \frac{C(\mathbb{U}_F(t))}{C(z\mathbb{U}_F(t))} - z^{1/\gamma_2} \right) \frac{d\overline{F}(z\mathbb{U}_F(t))}{\overline{F}(\mathbb{U}_F(t))}$$

and

$$D_3(s; t) := -x^{-1/\gamma_2} \int_x^\infty z^{1/\gamma_2} d \left( \frac{\overline{F}(z\mathbb{U}_F(t))}{\overline{F}(\mathbb{U}_F(t))} - z^{-1/\gamma} \right).$$

By applying Proposition B.1.10 in [de Haan and Ferreira \(2006\)](#) to both  $C$  and  $\overline{F}$  with integrations by parts, it is easy to verify that

$$|t\Lambda(x\mathbb{U}_F(t))C(x\mathbb{U}_F(t)) - (\gamma_1/\gamma)x^{-1/\gamma}| \leq \epsilon x^{-1/\gamma \pm \epsilon}.$$

Observe now that  $t\Lambda(x\mathbb{U}_F(t))C(x\mathbb{U}_F(t)) - (\gamma/\gamma_1)x^{1/\gamma}$  is equal to

$$((t\Lambda(x\mathbb{U}_F(t))C(x\mathbb{U}_F(t)))^{-1})^{-1} - ((\gamma_1/\gamma)x^{-1/\gamma})^{-1}.$$

By using the mean value theorem, the latter equals

$$\frac{(\gamma_1/\gamma)x^{-1/\gamma} - t\Lambda(x\mathbb{U}_F(t))C(x\mathbb{U}_F(t))}{(\psi(x; t))^2},$$

where  $\psi(x; t)$  is between  $(\gamma_1/\gamma)x^{-1/\gamma}$  and  $t\Lambda(x\mathbb{U}_F(t))C(x\mathbb{U}_F(t))$ . In view of the representation (6.27) and Potter's inequalities, applied to  $C$  and  $\overline{F}$ , with an integration by parts, we get  $t\Lambda(x\mathbb{U}_F(t))C(x\mathbb{U}_F(t)) \geq (1 - \epsilon)x^{-1/\gamma \pm \epsilon}$ . It follows that  $(\psi(x; t))^2 \geq (1 - \epsilon)^2 x^{-2/\gamma \pm 2\epsilon}$  and therefore

$$|t\Lambda(x\mathbb{U}_F(t))C(x\mathbb{U}_F(t)) - (\gamma/\gamma_1)x^{1/\gamma}| \leq (1 - \epsilon)^{-2} \epsilon x^{1/\gamma \pm \epsilon},$$

as sought. □